Document de Travail INED-MATISSE n° 2002-xx

Ressources Statistiques via

INTERNET

Un aperçu

Arnaud Bringé & Monique LE GUEN

INED CNRS-MATISSE

Atelier SAS
05 décembre 2002

Note
Ce document a été réalisé pour l'atelier SAS du 5 décembre 2002 à l'INED, dans le but de montrer les ressources variées apportées par Internet dans le domaine de la Statistique.
Ce document n'a rien d'exhaustif ! ! ! C'est un aperçu pour vous aider à démarrer.

Devant le gigantisme des ressources Statistiques présentes sur le Web, il est indispensable de structurer l'offre. Lors de cet atelier nous passerons succinctement en revue les différents types de ressources suivants, en les illustrant par des exemples : logiciels statistiques gratuits et téléchargeables, Extensions gratuites à des programmes existants, Documentations logicielles, Cours de Statistiques classiques en Français, Cours de Statistiques classiques en Anglais, Cours de Statistiques associés à SAS, Cours de Statistiques orientés Analyse Exploratoire des données (EDA), APPLETS, Glossaires de Statistiques.

A partir de ce document, il vous suffit de cliquer sur chaque adresse pour déclencher l'ouverture du site.

A vous de jouer…

bringe@ined.fr & leguen@univ-paris1.fr

Table des matières

L'environnement de travail

Les principaux navigateurs du marché :

Aujourd'hui, Internet Explorer domine très largement le marchés des browsers (Part de marché de 93% au 30 octobre 2001). Ceci peut en partie s'expliquer par le fait que le navigateur de Microsoft est fourni (et prêt à l'utilisation) en standard avec Windows. Netscape, actuellement disponible en version 7, a perdu une part importante de part de marchés, et n'occupe plus actuellement qu'une position marginale.

Adresses Web :

Internet Explorer 6 :
http://download.microsoft.com/download/IE60/fnlrtw/ie6/W98NT42KMe/FR/ie6setup.exe

Netscape 7 :
http://www.netscape.fr/telechargement/

Deux autres navigateurs, d'utilisation marginale existent :

- Mozilla, version libre de Netscape
  http://www.mozilla.org

- Opera : http://www.opera.com

Les plug-ins

L'utilisation d'un navigateur peut ne pas être suffisa nte pour profiter pleinement du contenu de certains sites web qui offrent, en plus du texte, des sons, animations et autres vidéos. Il est alors nécessaire de télécharger un programme additif qui va venir compléter les fonctionnalités de votre navigateur :

Si le plug-in est installé, il sera lancé automatiquement pour lire le fichier correspondant. Sinon le navigateur demandera l'application nécessaire pour le lire, en proposant de télécharger le plug-in adéquat pour lire le fichier inconnu. Les plug-in ainsi téléchargés seront dans la plupart des cas des fichiers compressées auto-installables.

Quelques exemples de plug-ins

- **Apple-Quicktime**

Adresse de téléchargement: http://www.apple.com/quicktime/download

- **Adobe Acrobat Reader**

Adresse de téléchargement : http://www.adobe.com/products/acrobat/readstep2.html

Permet de lire les fichiers au format PDF

- **Adobe SVG Viewer**

Adresse de téléchargement : http://www.adobe.com/svg/viewer/install/

Le langage SVG, basé sur XML, est en train de s'imposer comme un standard pour l'affichage de graphiques interactifs.

- **Real-Player**

Adresse de téléchargement : http://france.real.com/freeplayer_r1p.html

Permet de visionner un grand nombre de vidéos et autres sons disponibles sur le web. Permet notamment de lire les fichiers aux formats RealAudio.

- **Macromedia Flash Player et Shockwave player**

Adresse de téléchargement : http://www.macromedia.com/fr/downloads

Les moteurs de recherche

Un moteur de recherche est nécessairte pour rechercher une information. On distinguera :

Les moteurs de recherches simples

A tout seigneur tout honneur, le moteur le plus utilisé est google, www.google.fr, qui peut s'agrémenter d'une barre de menus sous Internet Explorer, instalable à partir de http://toolbar.google.com/intl/fr/ Ce site peut rechercher une image ou un mail posté dans un forum.

On peut aussi citer Yahoo (http://fr.yahoo.com/), AlltheWeb, (http://www.alltheweb.com/),

Les méta-moteurs

Ces outils permettent d'interroger simultanément plusieurs moteurs de recherche. Si la demande est trop imprécise, le nombre de réponses sera très important.

Vivisimo est un métamoteur (www.vivisimo.com), et a l'avantage de présenter les résultats par catégorie.

SearchOnline est un autre métamoteur extrêmement puissant, qui recherche simultanément sur 150 moteurs de recherche.

On peut aussi citer Metacrawler (www.metacrawler.com).

Les moteurs avec module de cartographie

Le moteur Kartoo (www.kartoo.com) affiche les résultats sous forme d'une carte, permettant de visualiser plus facilement et affiner les critères de recherches.

Quelles types de ressources statistiques ?

Devant le gigantisme des ressources présentes sur le Web, on peut dans un premier temps s'interroger sur les types de ressources disponibles. Nous passons ici succintement en revue ces différents types de ce que nous appelons ressources, en les illustrant brièvement par un exemple.

Les ressources que l'on peut trouver au niveau statistique sont de nature très diverses :

Des documentations logicielles

Le mode d'emploi d'utilisation d'un logiciel :
http://www.indiana.edu/~statmath/stat/sas/win/index.html

Des exemples de programmes, pour un logiciel déterminé :
http://www.sashelp.com/article.php?sid=50

L'application d'une méthode statistique, avec un ou plusieurs logiciels :
http://www.math.yorku.ca/SCS/Courses/grcat/grc6.html (Régression logistique avec Sas)

Des logiciels indépendants en téléchargements

D'utilisation gratuite et immédiate (freeware) :
http://cran.r-project.org/

D'utilisation payante, le plus souvent après une période d'essai (shareware)
http://www.xlstat.com/download.htm

Des extensions gratuites à des programmes existants

Des macros SAS :
http://www.sashelp.com/download.php?op=MostPopular

Une assistance à l'enseignement d'une méthode statistique

Cours de statistiques : http://www.lsp.ups-tlse.fr/Besse/Hyper/modlinhtml/modlinhtml.html (P. Besse, Univ. Toulouse)
Démonstrations interactives de principes statistiques (utilisation d'applets) :
http://www.pa.uky.edu/~phy211/graph_applets/plot_graph.html

Des "Espaces de vie" relatifs

Listes de discussions : http://www.listserv.uga.edu/archives/sas-l.html (Archives liste SAS)
Journaux en ligne : http://www.amstat.org/publications/jse/ (Journal of Statistics Education)

Les logiciels libres

Qu'entend-on par logiciel libre ?

En 1984, un chercheur du MIT, R. Stallman, pose les bases du logiciel libre, en pronant un développement logiciel basé sur la mise à disposition des codes sources. Le modèle économique de Stallman se fonde sur la licence GPL (Gnu Public Licence) qui donne obligation de mettre aussi à disposition de tout le monde toute amélioration pouvant être faite à partir de ces développements. Ce processus économique, à l'encontre des grands développements propriétaires sera accéléré avec l'avènement de Linux qui assure un noyau et un système d'exploitation à un environnement logiciel.

Un Logiciel est dit Libre s'il respecte les quatre règles suivantes

- Pouvoir être exécuté pour n'importe quel but;
- Pouvoir étudier son fonctionnement et l'adapter à ses besoins;
- Pouvoir redistribuer des copies;
- Pouvoir l'améliorer et diffuser les améliorations au public pour en faire bénéficier toute la communauté

Une tentative a été faite par l'OSI (Open Source Initiative) pour généraliser quelques fondements mis en œuvre par le logiciel libre, et assurer son développement économique, mais ceci a abouti à une fracture entre les logiciels libres (projets GNU de la FSF), et les logiciels Open Source (promus par l'OSI) permettant de ne redistribuer qu'une partie de code source. La querelle sur les interfaces KDE etGnome dans le développement de Linux est un exemple de cette querelle.

Tout savoir sur le logiciel libre : http://www.fsfeurope.org/index.fr.html
Tout savoir sur l'OSI : http://www.opensource.org/

Liste de logiciels libres GPL (projet GNU) :

- Liste des projets principaux : http://www.gnu.org/projects/projects.fr.html
- Liste des logiciels : http://www.gnu.org/directory/
- Liste des logiciels statistiques : http://www.gnu.org/directory/math/stats/

Logiciels libres pour l'enseignement : http://logiciels-libres-cndp.ac-versailles.fr/index.html

Le site http://sourceforge.net permet de trouver les logiciels en cours de développement. Ce site offre un support pour un travail coopératif dans le développement logiciel, base du fonctionnement du logiciel libre.

<u>Au niveau statistique, principaux logiciels libres dans le cadre du GNU :</u>

R : Basé sur le noyau S, développé à la Bell Corporation, ce clone du logiciel Splus, se caractérise par un développement actuel rapide, et par la prise en compte de très nombreuses méthodes statistiques. C'est un logiciel "command line", doté d'un puissant langage de programmation. Il se caractérise par une programmation orientée objet, et par un traitement matriciel des données.

Site Web : http://cran.r-project.org/

Un autre logiciel, **Gretl**, spécialisé dans les régressions et séries temporelles est actuellement en cours de développement.

Site Web : http://gretl.sourceforge.net/

Le logiciel **Tda**, permet notamment le traitement des analyses biographiques.

Site Web : http://www.stat.ruhr-uni-bochum.de/tda.html

Autres freewares :

D'autres freewares sont disponibles, mais ne figurent pas dans le contexte Gnu.

Cspro : Conception et traitement d'enquêtes. Logiciel conçu par les Nations Unies.

Adresse Web : http://www.census.gov/ipc/www/cspro/index.html

Epiinfo : Le Cdc, agence du département de la santé américain, a conçu et développé ce logiciel, utilisable non seulement dans le domaine médical, mais aussi pour le traitement de toute enquête. Contient un module de cartographie.

Adresse Web : http://www.cdc.gov/epiinfo/

Le site SciViews : http://www.sciviews.org/

Les applets : Première approche

Définition : Une applet est une application Java, qui s'exécute dans un navigateur Web.
Programme en langage Java qui s'exécute dans le navigateur du client, ce qui permet notamment une interactivité

Panorama d'applets disponibles

Les sites SEL (Statistiques En Ligne) et SMEL (Statistiques Médicales En Ligne) proposent un nombre important d'applets afin de sensibiliser aux notions élémentaires de la statistique. L'interactivité est complète et permet de mieux comprendre certaines notions. Ces sites contiennent en outre un lexique, une sélection d'articles, et un cours de statistiques.

Adresses Web: http://www.inrialpes.fr/sel/ et http://www.math-info.univ-paris5.fr/smel/

Tracé de courbe : http://www.pa.uky.edu/~phy211/graph_applets/plot_graph.html

Multiples fonctionnalités statistiques : http://www.stat.sc.edu/webstat/

Extension du logiciel R en Rweb

Le logiciel R peut être exécuté à distance via une application javascript. Un programme est exécuté sur un serveur puis le résultat retourné dans une fenêtre séparée. Le logiciel n'est pas installé en local !

Sas

Site sas.com

Le site Sas.com est le portail de toutes les informations relatives à Sas-France. Du fait de sa structure, il est extrêmement difficile de rechercher une information. La partie support est extrêmement riche (http://www.sas.com/service/index.html)

Le site Sas France www.sas.com/france reprend une structure équivalente. A noter, un ensemble de FAQ (http://www.sas.com/offices/europe/france/services/techsup/Question.html)

Le site Sas Academic est accessible à l'adresse http://www.sas.com/offices/europe/france/academic/index.html, mais on y notera la pauvreté des informations, et l'absence de forums !

La documentation SAS de la version 8 (Online Doc) est disponible à l'adresse http://v8doc.sas.com/sashtml/ (Pour une 1ère utilisation, cliquer sur annuler).

Un Support technique est disponible par email. On indique son numéro de site, le système et le module objets de la question, et un texte de question. Un fichier log peut être joint.
Adresse Web :
http://www.sas.com/offices/europe/france/services/techsup/submit_emits3.html

Des macros statistiques et exemples utilisation procédures statistiques sont disponibles sur le site SAS :
http://ftp.sas.com/techsup/download/stat/

Il est à noter que les actes des réunions Sugi 1997 à 2002 sont en ligne. Il est possible de télécharger les supports des présentations au format PDF. Ces documents témoignent de la richesse et de la grande qualité des présentations.

Adresse Web : http://www.sas.com/usergroups/sugi/proceedings/index.html

Un sourire :
Enfin un support technique sur la version 6.12 de Sas !
http://www.sas.com/service/techsup/intro.html

Les sites ressources

D'autres sites fournissent des ressources, documentations ou macros. Ainsi le site de Michaël Friendly, fournit de nombreuses informations sur les graphiques statistiques :

Paper : http://www.math.yorku.ca/SCS/sugi/sugi16-paper.html
Macros : http://www.math.yorku.ca/SCS/sasmac/

Un certain nombre de trucs et astuces en Sas sont disponibles sur ces sites :
http://tips.pwcons.com/
http://www.sashelp.com/

Un manuel d'introduction à Sas est disponible à cette adresse :
http://www.indiana.edu/~statmath/stat/sas/win/index.html

FAQ sur l'ODS : http://www.sas.com/rnd/base/topics/odsprinter/faq.html

Archives de la liste SAS-L : http://www.listserv.uga.edu/cgi-bin/wa?S1=sas-l
Cette adresse permet de retrouver les messages relatifs à un mot clé, sujet ou auteur …

Jeux et divertissements statistiques

Pour tester ses connaissances :
http://www.economics.pomona.edu/StatSite/statgame.html
http://www.economics.pomona.edu/StatSite/Quiz.html

Professor Mean : http://www.cmh.edu/stats/

Applications

Recensement français : http://www.recensement.insee.fr/
Recensement US : http://factfinder.census.gov/servlet/BasicFactsServlet
Pyramide des âges (Dos) : http://www.ined.fr/recherches/projetsweb/index.html

PARTIE II : Ressources WEB pour l'Apprentissage et l'Enseignement de la Statistique

Introduction

Côté Cours
Le WEB permet d'accéder en temps réel, puis d'imprimer, de télécharger, de copier des cours de statistiques classiques, (notes ou cours complets), présentés de manière séquentielle comme un document papier.
A côté de cette forme classique, une nouvelle forme de cours dits cours MULTIMEDIA se développe. Ces cours utilisent les possibilités du WEB, documents en Hypertexte (.html) avec des liens vers d'autres documents et des exécutions d'applications APPLETS permettant de réaliser des calculs ou des représentations graphiques statiques ou dynamiques, en ligne.

Ces cours font appel à de nombreuses ressources qui sont éparpillées sur tout le WEB. C'est une intégration **phénoménale** qui permet l'accès à des documents de plus en plus pédagogiques.

Côté Logiciels
Certains centres américains d'enseignement de la Statistique via le WEB ont développé en parallèle à ces sources documentaires, des applications-logiciels essentiellement basées sur l'Analyse Exploratoire des données. Ces projets s'inspirent des travaux de J. W. TUKEY. Citons VELLEMAN à l'Université de Cornell, BEHRENS et GLASS à l' Université de l'Arizona (ASU), YOUNG à l'Université de Nord Carolina (UNC). Ces centres sont des ressources aussi bien pour les élèves (**Learning**= Apprentissage) que pour les enseignants (**Teaching**=Enseignement).

LISTES de liens

Six portails importants vers des ressources pour l'apprentissage et l'enseignement de la statistique :

- J Puranen Université de Helsinski http://noppa5.pc.helsinki.fi/links.html 34 pages de liens ! ! ! ! Un point d'entrée spécifique à l'enseignement de la Statistique : http://noppa5.pc.helsinki.fi/papers/paris/index.html
- R. LOCK de St Lawrence University Canton NY (STLAWU) http://it.stlawu.edu/~rlock/tise98
- FRIENDLY Université de Toronto : http://www.math.yorku.ca/SCS/StatResource.html
- http://home.okstate.edu/homepages.nsf/toc/onlinetutorialsstatistics
- http://www.stat.ucl.ac.be/ISpersonnel/lecoutre/stats/
- http://members.aol.com/johnp71/javastat.html#Descriptives

Le CTI : un portail sur presque tous ! les logiciels Anglais de Statistiques . On y trouve Sphinx mais pas SPAD

- http://www.stats.gla.ac.uk/cti/links_stats/software.html

COURS de Statistiques en FRANCAIS

Cours Classique de Stat Des INRIA et Université René Descartes Paris V

Projet **SEL** : Statistique en Ligne avec des Applets de simulations
http://www.inrialpes.fr/sel/index.html
et projet **SMEL** : Statistiques Médicales En Ligne
http://metice.univ-montp3.fr/~miap/sel/

Les projets SEL et SMEL propose une initiation interactive à la statistique, articulée en trois couches.

- La couche ARTICLES propose des textes, contenant des exemples d'utilisation de la statistique.
- La couche LEXIQUE contient un index des termes statistiques référencés dans les articles. Certains termes sont illustrés par des APPLETS. (cf. rubrique **Simulations et Diagrammes en bâtons…**).
- La couche COURS est un cours de statistique au sens classique.

Cours multimedia de Stat Des CNAM

Cours de Statistique Descriptive divisé en 5 leçons, avec des présentations, développements, résumés, exercices interactifs, lexique.
Déchargement à distance des cours, des exercices, des fichiers de données ou des utilitaires de statistique...
Liens vers d'autres sites.
ST@TNET http://www.cnam.agropolis.fr/auf/

Remarque: La navigation dans ce cours est très facile.

CICT et Laboratoire de Statistique et Probabilités de l'Université Paul Sabatier de Toulouse

BESSE
Cours de Statistique Descriptive Uni, Bidimensionnelle,
Cours de Statistique Descriptive Multidimensionnelle : Techniques factorielles de base
Cours sur le Datamining
Initiation à R etc…
http://www-sv.cict.fr/lsp/Besse/enseignement.html

SAINT PIERRE -http://www.cict.fr/cict/personnel/stpierre
Exposés de J. Saint-Pierre : http://www.cict.fr/cict/personnel/stpierre/exposes.html
Notes de cours : Éléments primaires de la Modélisation en Traitement de Données
http://www.cict.fr/cict/personnel/stpierre/expose-27-02-98/expose.html
Description de Proc SAS : Proc CALIS de SAS http://www.cict.fr/cict/personnel/stpierre/sas-calis/sas-calis.html

➔**TENENHAUS HEC**
Cours de Statistique Descriptive et Analyse de données, présentation d'un dendogramme animé (très pédagogique, aide à la compréhension de la classification) (version en français et anglais).
http://Siad.hec.fr/statistiques/tenenhaus/Default.htm

C. GRASLAND, Université de Paris 7 :
Introduction à la statistique en sciences sociales :
http://ibm2.cicrp.jussieu.fr/grasland/STAT98/STAT98.htm

INRA
Développement de formations à la statistique, avec supports en ligne.
http://www.inra.fr/bia/T/schnu/FPstat/FP2.html
Exemple :
Algèbre matricielle : http://www.inra.fr/bia/T/schnu/FPstat/moduleA-v1.html

Ressources humaines Canada :
Cours sur les enquêtes (cueillette !) :
http://www.qc.hrdc-drhc.gc.ca/socio-ec/enq_stat/default.html#M%E9thodes%20de%20collecte

Autres cours de statistiques :

Statistique et analyse de données en gestion :
C. Benavant : http://christophe.benavent.free.fr/cours/stat/stat.htm

Statistiques descriptives pour IAE :
D. Grau : http://wwwbay.univ-pau.fr/~grau/1A/intro.html

Techniques d'analyse quantitative en psychologie :
J. Begin (Canada) : http://www.er.uqam.ca/nobel/r30574/PSY1300/

Cours complet de statistiques descriptives, avec tables usuelles (Student, Normale, Chi2) :
Biostatistique (Paris6) : http://www.chups.jussieu.fr/polys/biostats/poly/POLY.TDM.html

Méthodes statistiques :
Statistique Non Paramétrique : http://www.math.niu.edu/NPAR/

Cours et Documentations SAS® en Français

CICT et Laboratoire de Statistique et Probabilités de l'Université Paul Sabatier de Toulouse

BESSE
Macros ACP, AFD (Proc Candisc), Bootstrap, CAH, Nuées Dynamiques etc…
http://www.lsp.ups-tlse.fr/Besse/pub/sas/

Institut de Statistique de Université Catholique de Louvain

Vue d'ensemble http://wwwsas.stat.ucl.ac.be/sasdiscute.html
Didacticiels en SAS : http://wwwsas.stat.ucl.ac.be/sasdiscute/doc/tutorials.html

http://sas.stat.ucl.ac.be/sasdiscute/doc/tutorials/base/data.html

Programmes exemples SAS/BASE :
http://wwwsas.stat.ucl.ac.be/sasdiscute/doc/samples/base/samp_base.html

Université de Montreal

Département de Statistique http://www.dms.umontreal.ca/~stat/
Logiciel SAS : http://www.DMS.UMontreal.CA/~stat/logiciels/sas/
Mini guide : http://www.DMS.UMontreal.CA/~stat/logiciels/sas/mini-guide/SAS.html

Des informations sur des logiciels de Statistique http://www.dms.umontreal.ca/~stat/#log
COURS de SAS de base complet généralités et procédures
http://www.dms.umontreal.ca/~stat/sas/SAS/SAS.html

Université de Laval (Ca)

Guide d'utilisation de SAS http://www.ecn.ulaval.ca/aide/sas/
Régression logistique : http://w3.res.ulaval.ca/cours-epm-64312/Default.htm

Cours et Documentations SAS® en Anglais

Université de WISCONSIN : http://www.stat.wisc.edu/computing/sas/
Université de TORONTO : http://www.yorku.ca/dept/psych/lab/sas/index.htm/

Université de MICHIGAN http://www.umich.edu/
Permet d'accéder à : CSCAR-Center for Statistical Consultation and Research

Cours de SAS http://www.umich.edu/~cscar/software/sas.html
Langage, Fonctions, Recodages en SAS http://www.umich.edu/~cscar/sas/sasrules.html
Graphiques en SAS http://www.umich.edu/~cscar/sas/grafcat.html
Exemples en SAS http://www.umich.edu/~cscar/sas/code.html
Etc…
Par ex : Comment supprimer les sauts de pages en SAS

http://www.umich.edu/~cscar/sas/sascode/nopagebreak.sas

Université de CARNEGIE MELLON

STATLIB : est un système qui distribue des logiciels de Statistiques, des données et des informations électroniques (par mail, FTP, WWW). « *Data, Software and News from the Statistics Community*"

CARNEGIE STATLIB pour s'enregistrer (gratuit Pass leguen/1efbkjlA)
http://lib.stat.cmu.edu/
Choisir menu : ➔CMU statistics puis ➔Technical report ➔Articles en ligne exemple :
http://www.stat.cmu.edu/www/cmu-stats/tr/tr769/tr769.html
STATLIB---pour accéder à des données ➔ Datasets http://statmib@lib.stat.cmu.edu/datasets/
AGRESTI Données de son livre Categorical Data Analysis
http://statmib@lib.stat.cmu.edu/datasets/agresti
STATLIB pour rechercher des messages sur un sujet, selon l'auteur etc… :
http://lib.stat.cmu.edu/cgi-bin/iform?SNEWS

Exemple :

Search Terms will be combined by AND
Field: SUBJECT Term: regression
Field: SUBJECT Term: logistic
Field: FROM Term: Ripley

Données pour l'Enseignement

DASL via STATLIB http://statmib@lib.stat.cmu.edu/DASL/
DASL (pronouncé "dazzle") Data and Story Library

DASL est une bibliothèque en ligne contenant des données (**Data**) et des Histoires (**Story**) au sens de méthodes utilisables sur ces données.
Un moteur de recherche efficace permet de trouver des données classées par thèmes (**Topics**) et par méthodes (**Methods**), utile pour les enseignants comme pour les élèves.
Exemples :
➔ choisir Topics (rubriques) : Economics/ Women in the Labor Force Story/ …
http://lib.stat.cmu.edu/DASL/Stories/WomenintheLaborForce.html
➔ cliquer Methods: Boxplot, donne la liste des topics traitant de boxplots
http://lib.stat.cmu.edu/DASL/Stories/DraftLottery.html

Cours de Statistiques orientés EDA

Bibliographie de John W. Tukey

David R. Brillinger , (2001) “John W. Tukey *: The life and Professional Contributions*”
http://stat-www.berkeley.edu/users/brill/life19.pdf

Memories of J. W. Tukey: http://stat.bell-labs.com/who/tukey/index.html
http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Tukey.html

Remonter au niveau Mathematicians pour trouver d’autres bibliographies de mathématiciens.

HORBER Eugène et MIRAGE

Le site WEB de Eugène HORBER à Genève et celui de l’Association MIRAGE (Mouvement International pour le développement de la Recherche en Analyse Graphique et Exploratoire) ont plusieurs liens vers d’autres sites orientés Analyse Exploratoire des Données (EDA en anglais).
Ces deux sites sont des points de départ vers la plus part des sites cités ci-après.

HORBER EDA http://www.unige.ch/ses/sococ/eda/
MIRAGE http://www.unige.ch/ses/sococ/mirage/

- Liste sur l’Exploration et la Visualisation
  http://www.unige.ch/ses/sococ/eda/explor/expsoft.html

Remarque
Un cours de statistiques orienté EDA est essentiellement dépendant d’un logiciel, donc cours et logiciels sont imbriqués.

VISUALISATION

Transparents avec des graphiques sur, les graphiques, la perception graphique, les Co-plots, les graphiques multivariés, les graphiques dynamiques etc…
http://maddog.fammed.wisc.edu/course/lecture1
http://maddog.fammed.wisc.edu/course/lecture2

VELLEMAN Paul Cornell Université et Data Description

PAUL F. VELLEMAN est professeur de Statistiques et ancien élève de TUKEY. Dès 1985 il développe sur Macintosh à l'Université de Cornell un logiciel **Datadesk** pour l'enseignement de la Statistique à partir de l'Analyse Exploratoire des données

DATADESK de Velleman

http://www.datadesk.com/DataDesk/
pr télécharger une version démo :
http://www.datadesk.com/DataDesk/download.shtml

ActivStats

http://www.datadesk.com/ActivStats
Cours de statistiques multimedia développé par Data Description, société créée par Velleman pour développer et commercialiser ses produits (Datadesk à l'origine).
Activstats tourne sur Mac et PC en version étudiant et professionnel.

ActivStats Interactive Tools
http://www.datadesk.com/ActivStats/toys/conf.shtml



ActivStats Interactive Tools
http://www.datadesk.com/ActivStats/toys/hyptest.shtml
ActivStats Interactive Tools
http://www.datadesk.com/ActivStats/toys/central.shtml
ActivStats Interactive Tools
http://www.datadesk.com/ActivStats/toys/small.shtml

BEHRENS John- Arizona State University- ASU et SIIP

http://research.ed.asu.edu

Un cours EDA : Graphical and Exploratory Data Analysis
http://research.ed.asu.edu/classes/eda/topics.html cliquer sur les différentes semaines (Week)

Pour aller vers http://research.ed.asu.edu/classes/edp552/behrens/class1.html ou class2.html etc.
une présentation très claire des Boxplots.
http://research.ed.asu.edu/siip/briefs/boxplots.computing.html
Une collection de graphiques **Data Gallery** développé avec **SIIP Statistical Instruction Internet Palette** http://research.ed.asu.edu/siip/

Dr. B's Wide World of Web Data
C'est un cours de Statistique crée en 1994.
http://research.ed.asu.edu/siip/webdata/ qui intègre le produit SIIP.

Article de réflexion philosophique et méthodologique à partir des idées du philosophe anglais PEIRCE (1839-1914) présenté par Yu, Chong Ho, « *Abduction ? deduction? Induction? Is there a logic of Exploratory Data Analysis ?»* Annual Meeting of American Educational Research Association, New Orleans Louisiana, April, 1994.
http://seamonkey.ed.asu.edu/~behrens/asu/reports/Peirce/Logic_of_EDA.html

GLASS Gene de l' ASU

Cours de Statistique http://glass.ed.asu.edu/stats/
Synopsis de la Statistique http://glass.ed.asu.edu/stats/lesson0
Tendance variabilité Skewness : http://glass.ed.asu.edu/stats/lesson2/index.html#f
Calcul de Stat Descriptives en ligne : http://glass.ed.asu.edu/cgi-bin/basic.pl

YOUNG Forrest de l'UNC et son logiciel VISTA

Professeur de Psychométrie de l'Université de South Carolina depuis 30 ans.
Fut consultant pour SAS, BMDP, S-PLUS. Il a écrit des modules pour SAS, SPSS.
Ses recherches portent sur « *Multidimensional and Nonlinear Multivariate Data Analysis*»
Auteur des produits ALSCAL, ViSta et de la Proc MDS de SAS.

Site de YOUNG F. http://forrest.psych.unc.edu/
Utiliser ensuite les menus à gauche pour accéder à sa Biblio, Enseignement, VISTA etc.…

Enseignement (**Teaching**)
Puis ➔ **Intro Psy Stat** puis ➔**Schedule** et ➔**lecture note**
ou directement http://forrest.psych.unc.edu/teaching/p30/index.html

« *In this course, the introduction to statistics and data analysis is presented with an unusual emphasis: The goal is to "**See what your data seem to say**"* (TUKEY, Exploratory Data Analysis) »

On Significance Testing Paradigm pointe sur un article écrit par LYLE V. JONES and JOHN W. TUKEY le 11/10/99 « *A Sensible Reformulation of the Significance Test* »

Cours de Computing et de Statistique destiné à des statisticiens désirant développer des procédures de statistiques et de Visualisation
http://forrest.psych.unc.edu/teaching/p285/index.html

ViSta : The Visual Statistics System

Système de visualisation Statistique écrit en LISP (10 ans de développement) par YOUNG F.

ViSta - http://forrest.psych.unc.edu/research/ clic Spread

LANE David - Arizona Stat Univ-

Rice Virtual Lab in Statistics

Cours, Java Applets pour démontrer des concepts Statistiques (**Simulations/Demonstrations**), Labo d'Analyse (**Analysis lab**), liens vers d'autres ressources sur le Web,
http://www.ruf.rice.edu/~lane/rvls.html essayer les menus.

HYPERSTAT

Cours de STAT en ligne de LANE DAVID
http://davidmlane.com/hyperstat/index.html
Pour sourire : sélectionner dans la page le menu en bas à gauche « **Statistics?! Don't Make Me Laugh!** »

FRIENDLY Université de York Toronto

Spécialiste des représentations graphiques pour données catégorielles (comme le *Mosaic-plot* étendu pour des tableaux croisés à n dimensions) .

Site : http://www.math.yorku.ca/SCS/friendly.html
à voir entre-autres, dans la mouvance de Seymour Papert: *SAS/GRAPH Meets the Logo Turtle* http://www.math.yorku.ca/SCS/sugi/saslogo.html
Article : Visualizing Categorical Data
http://hotspur.psych.yorku.ca/SCS/Papers/casm/casm.html
Article : Les graphiques statistiques sous Sas :
http://www.math.yorku.ca/SCS/sugi/sugi16-paper.html/

STATLAB Yale University (HARTIGAN)

Cours de STAT et de SAS http://statlab.stat.yale.edu chercher HARTIGAN – HARTIGAN est un des développeurs des mosaic- plots

NIST/ Engineering Statistics HandBook et DATAPLOT

http://www.itl.nist.gov/div898/handbook/index.htm
➔ Choisir rubrique Explore
Exemple de Anscombe : http://www.itl.nist.gov/div898/handbook/eda/section1/eda16.htm

Logiciel graphique gratuit : DataPlot
http://www.itl.nist.gov/div898/software/dataplot/ ➔ choisir Summary puis Graphic Gallery
Un cours complet de Statistiques orienté EDA et destiné à des Ingénieurs.

Exploring DATA

http://exploringdata.cqu.edu.au/
Ce site, très ergonomique fournit des ressources matérielles pour les enseignants qui enseignent les bases de la Statistique dans la mouvance EDA.

Sites orientés Education avec APPLETS

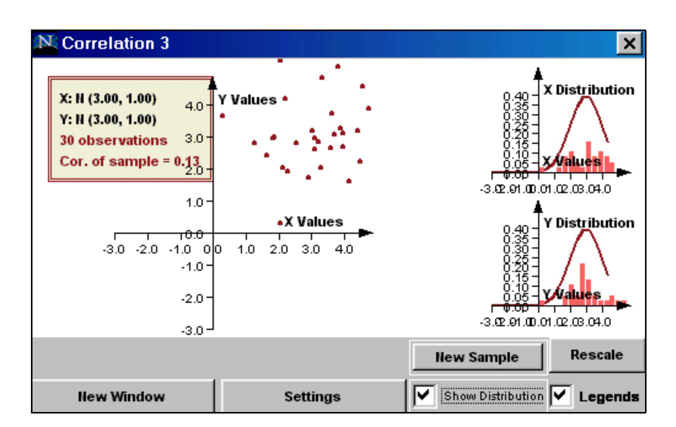
Les sites cités ci-dessous ont été développés dans le but de faciliter l'apprentissage des concepts de la Statistique par le biais des visualisations et des simulations.
Si un site propose plusieurs applets, seul un ou deux exemples sont cités.

La Planche de Galton

La simulation de la **planche de Galton** est accessible sur les sites suivants :
http://acad.cgu.edu/wise/appletsf.shtml , cliquez ensuite sur l'Applets Galton's
Quincunx de Galton sur le site de Berrie http://www.stattucino.com/berrie/dsl/Galton.html
The Central limit in Action (la plus spectaculaire) :
http://www.rand.org/methodology/stat/applets/clt.html

VESTAC Visualization and Experimentation with STAtistical Concepts

de l'Université Catholique de LOUVAIN http://www.kuleuven.ac.be/ucs/java/
Ce site est l'un des plus complets. Il concerne les rubriques : Basic, Tests, Regression, ANOVA
Exemple : Applet sur la Corrélation dans Basic



Voir également dans **Basic** une simulation du coefficient de corrélation de la loi normale bivariée.

WISE Web Interface for Statistics Education

Ce site concerne les distributions, la corrélation, la régression, ANOVA
http://acad.cgu.edu/wise/appletsf.shtml
Voir dans Regression la visualisation de la décomposition de SS Total=SS Model + SS Error
Le fonctionnement et l'objet de chaque applet de WISE sont très bien documentés.

Berrie

Applets sur les distributions normale, Binomiale, Poisson etc. sur le site de Berrie
http://www.stattucino.com/berrie/index.html

Darmouth

Site Education de Darmouth :
http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/applets.html
Ce site a de nombreux lien vers d'autres sites éducatifs. Voir la non robustesse d'un histogramme en changeant dynamiquement les tranches de l'histogramme : http://www.stat.sc.edu/~west/javahtml/Histogram.html

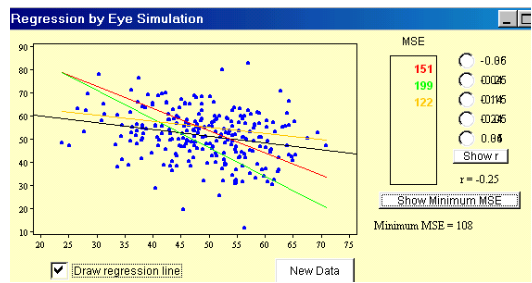
Comment la moyenne d'échantillon tend vers la moyenne de la population avec Tirage selon des lois de probabilités (Gauss, Exponentielle, Uniforme) :
http://arbitrage.byu.edu/sample.html

Lancement de dés pour simuler le Théorème Central limite :
http://www.stat.sc.edu/~west/javahtml/CLT.html

Visualiser les intervalles de confiance de la moyenne (tirage d'échantillons de taille 5 dans une loi N(0,1)). http://www.stat.sc.edu/~west/javahtml/ConfidenceInterval.html

RICE Rice Virtual Lab in Statistics

Ce Site développé par Davis M. LANE permet d'accéder à des cours et des applets.
COURS et APPLETS http://www.ruf.rice.edu/~lane/rvls.html
Index : http://www.ruf.rice.edu/~lane/stat_sim/index.html
Histogramme et Estimateur de densité par la méthode du noyau (kernel)
http://www.ruf.rice.edu/~lane/stat_sim/histogram/index.html
« Regression by eye » pour estimer la droite de régression:
http://www.ruf.rice.edu/~lane/stat_sim/reg_by_eye/index.html



APPLETS sur l'effet de l'ajout d'un point dans une régression
http://www.stat.sc.edu/~west/javahtml/Regression.html

Perception graphique selon les échelles en X et Y :
http://noppa5.pc.helsinki.fi/koe/aratio/aratio1.html

Cliquer sur **Clic Data** Pour voir les données, cliquer sur Sulje pour fermer la fenêtre !!!

TISE ***T**echnology **I**n **S**tatistics **E**ducation*

http://it.stlawu.edu/~rlock/tise98
RLOCK et TISE Choisir *OUT Line for the Talk* et voir les démonstrations Java (Java Demonstration)

Cours : http://it.stlawu.edu/~rlock/tise98/courses.html
Démonstration Java du coefficient de CORRélation :
http://it.stlawu.edu/~rlock/tise98/java.html

TISE mentionne les liens vers :
Essayer de deviner la corrélation avec JAVA CORR GUESS
http://www.stat.uiuc.edu/~stat100/java/GCApplet/GCAppletFrame.html
idem avec JAVA REG GUESS - Applet sur la Regression -
http://www.stat.sc.edu/~west/javahtml/Regression.html

GASP Globally Accessible Statistical Procedures

http://www.stat.sc.edu/rsrch/gasp/

Domaine des Réseaux de Neurones

présentation générale sur le Data Mining
http://www.statsoft.com/textbook/stdatmin.html

Logiciel gratuit sur Internet : SNNS Stuttgart Neural Network Simulator
http://rfhs8012.fh-regensburg.de/~saj39122/jfroehl/diplom/e-index.html
excellente introduction pour accéder à une simulation des cartes de Kohonen cliquer sur ➔ Applet Sample

Ce site donne des informations utiles pour les novices en réseaux de neurones :
ftp://ftp.sas.com/pub/neural/FAQ.html#A2

Commentaires de Warren Sarle de SAS sur les réseaux de neurones
http://www.pitt.edu/~wpilib/statfaq/neurfaq.html

IASE Journaux & Revues orientés Education

Association Internationale pour l'Enseignement de la Statistique (IASE)
http://www.cbs.nl/isi/iase.htm

L'Association Internationale pour l'Enseignement de la Statistique est la plus récente des sections de l'Institut International de Statistique (IIS). Elle a été créée en 1991. L'association publie une NEWSLETTER. Toutes les Newsletters de 1996 à 1999 sont disponibles sur le site http://www.ugr.es/local/batanero/

Journal of Statistics Education : http://www.amstat.org/publications/jse/
Journal of Statistical Software : http://www.stat.ucla.edu/journals/jss/

CHANCE

http://www.dartmouth.edu/~chance/
Equipe de statisticiens concernés par l'enseignement de la Statistique. Parmi ses membres quelques noms importants: J. Laurie Snell et Peter Doyle de Dartmouth College, Joan Garfield de l'University of Minnesota, Tom Moore de Grinnell College, etc…. Ils produisent des COURS et publient la Revue: **Chance News**.

Pour s'abonner à Chance News (gratuit) :
http://www.dartmouth.edu/~chance/chance_news/news.html

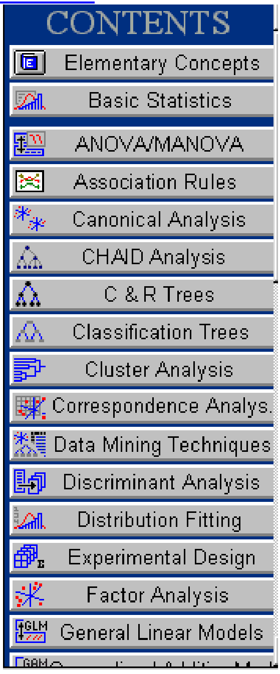
Autres associations de statisticiens

Société Française de Statistique (SFdS) : http://www.sfds.asso.fr
Association Penombre : http://www.unil.ch/penombre/
ISI, International Statistical Institute : http://www.cbs.nl/isi/
IASC, International Association of Statistical Computing : http://www.stat.unipg.it/iasc/
RSS, Royal Statistical Society : http://www.rss.org.uk
ASA, American Statistical Association : http://www.amstat.org/

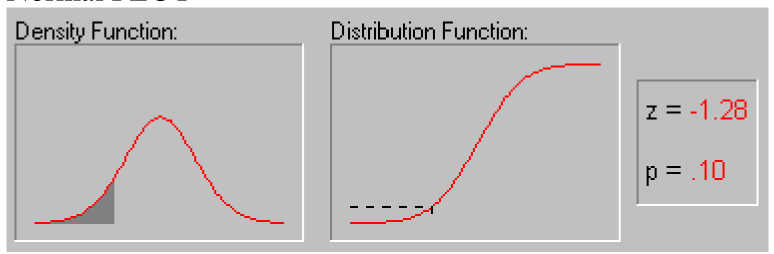
GLOSSAIRES de statistiques sur le Web

Statsoft -Très complet

Le dictionnaire électronique de Statsoft : très complet avec des images et graphiques animés
http://www.statsoft.com/textbook/stathome.html



Ex :
Basics Statistics Normal PLOT



Correspondance Analysis même « Benzecri » est cité !
"For a comprehensive description of this method, computational details, and its applications (in the English language), refer to the classic text by Greenacre (1984). These methods were originally developed primarily in France by **Jean-Paul Benzérci** in the early 1960's and 1970's (e.g., see **Benzérci, 1973;** see also **Lebart, Morineau, and Tabard,** 1977), but have only more recently gained increasing popularity in English-speaking countries (see, for example, Carrol, Green, and Schaffer, 1986; Hoffman and Franke, 1986)."

SportSciences

A new view of Statistics :
http://www.sportsci.org/resource/stats/index.html

Glossaire Eurostat conversion multilingues
http://europa.eu.int/comm/eurostat/research/index.htm ➔ Choisir Isi Glossary

STEPS - Plus léger
STEPS : S Tatistical **E**ducation **T**hrough **P**roblem **S**olving,
Définition des termes statistiques les plus courants avec entrée par alphabet
http://www.stats.gla.ac.uk/steps/glossary/index.html
Un exemple: qu'est ce qu'une statistique ?
http://www.stats.gla.ac.uk/steps/glossary/basic_definitions.html#stat

Glossaire de termes statistiques Statlets
http://www.statlets.com/usermanual/glossary.htm

STATISTICA - Manuel électronique
http://www.statsoftinc.com/textbook/glosfra.html

Bibliothèques de programmes et macros SAS

Macros statistiques : http://www.sas.com/service/techsup/faq/stat_macro.html
Site de Arnold Schick : http://members.tripod.com/~schick/
Site SasHelp : http://www.sashelp.com/
Macros graphiques de Friendly: http://hotspur.psych.yorku.ca/SCS/sasmac/index.html/
Site SFDS hébergé à l'INRIA : http://www-rocq.inria.fr/axis/Sfds_gls/MACROS/index.htm

Les jeux de données

Jeux de données validées pour applications statistiques : http://www.nist.gov/itl/div898/strd/

Centre Quetelet : http://www.centre.quetelet.cnrs.fr/

Statistiques de l'OCDE : http://www.oecd.org/statlist.htm

Statistiques démographiques : http://www.ined.fr

Accès à des liens sur des jeux de données : http://odwin.ucsd.edu/idata/

Recensement Insee 99 : : http://www.recensement.insee.fr/

Pour terminer :

Beaux Graphiques Historiques

LA QUINCONCE DE GALTON http://www.stat.ucla.edu/history/quincunx2.gif

PLAYFAIR : http://www.wmich.edu/ssc/about.html

MINARD site de TUFTE http://www.edwardtufte.com/337099670/tufte/minard

POSTER: NAPOLEON'S MARCH

Probably the best statistical graphic ever drawn, this map by Charles Joseph Minard portrays the losses suffered by Napoleon's army in the Russian campaign of 1812. Beginning at the Polish-Russian border, the thick band shows the size of the army at each position. The path of Napoleon's retreat from Moscow in the bitterly cold winter is depicted by the dark lower band, which is tied to temperature and time scales. Exquisitely printed in two colors on fine paper, 22" x 15".